JOURNAL OF DIFFERENTIAL EQUATIONS 88, 155–174 (1990)

The Most Likely Path of a Differential Inclusion

ALBERTO BRESSAN

Department of Mathematics, University of Colorado, Boulder, Colorado 80309-0426

Received June 7, 1989

1. INTRODUCTION

Multivalued differential equations of the form

$$\dot{x}(t) \in F(x(t)), \qquad t \in [0, T] \tag{1.1}$$

can be used, in alternative to stochastic processes, as a model for systems whose evolution is non-deterministic. Compared with the rich mathematical theory which is currently available for random processes, however, differential inclusions have remained a rather thin subject, concerned mainly with the existence of solutions and with the topological properties of the set of trajectories. The present paper is an attempt to widen the scope of this theory, introducing the notion of "likelihood" for a solution of a differential inclusion and formulating a corresponding class of estimation, prediction, and filtering problems.

To do this, a simple approach first comes to mind: define some probability measure on the set of trajectories of (1.1) and formulate the prediction and filtering problems accordingly. Unfortunately, there seems to be no canonical probability distribution which is supported precisely on the solution set of (1.1). The choice of any particular distribution would thus require considerable additional information about the system to be modeled. For this reason, we consider a new definition of "likelihood" which is entirely independent of probability theory, relying solely on the metric structure.

If $\mathscr{F}$ denotes the family of all solutions of (1.1) and $u \in \mathscr{F}$, a rough estimate of how many trajectories $v \in \mathscr{F}$ remain close to u is given by

$$\beta\{\dot{v} : v \in \mathscr{F}, |v(t) - u(t)| < \varepsilon \text{ for all } t \in [0, T]\}, \tag{1.2}$$

where β denotes the Hausdorff measure of non-compactness in $\mathscr{L}^2$ and $\varepsilon > 0$. As $\varepsilon \to 0$, the infimum of the quantities (1.2) is a well defined number,

0022-0396/90 $3.00

which we call the likelihood of u. Our main result shows that this number can be explicitly computed by an integral formula. For the Kuratowski measure of non-compactness of a decomposable subset of $\mathscr{L}^1$, a somewhat similar formula was proved in [2]. Our choice of the Hausdorff measure in $\mathscr{L}^2$ is largely motivated by the fact that, in this setting, the likelihood of a trajectory $u(\cdot)$ usually has a simple expression, depending only on the Chebyshev radius of $F(u(t))$ and on the distance between $\dot{u}(t)$ and the Chebyshev center of $F(u(t))$. The corresponding prediction and filtering problems can thus be written in the form of classical problems in the calculus of variations.

The possibility of defining the "most likely path" for a differential inclusion in a non-probabilistic context was first suggested by A. Cellina, whose remarks provided the initial motivation for this research. It is well known that several qualitative aspects of probability theory have a purely topological counterpart, formulated in terms of Baire Category. A comprehensive account of the analogies between measure and category can be found in Oxtoby's book [10]. The present paper yields an example of a quantitative aspect of probability theory which has a non-probabilistic counterpart, defined using measures of non-compactness.

Basic notations and definitions are collected in Section 2. The integral formula expressing the likelihood is stated in Section 3 and proved in Sections 4–6. In Section 7 we define the likelihood $L(\bar{x})$ for a point $\bar{x}$ to be approached by solutions of a multivalued Cauchy problem, and show that $L(\bar{x})$ is the value function of a corresponding optimization problem. The previous theory is then used in the last section, providing a rigorous mathematical formulation of a family of estimation, prediction, and filtering problems, in a context which is entirely independent of probability theory.

2. Preliminaries

In this paper we write $|\cdot|$ for the euclidean norm on $\mathbb{R}^n$, while $B(x, r)$, $\bar{B}(x, r)$ denote the open and the closed ball centered at x with radius r, respectively. We write $\bar{A}$ and $\overline{\text{co}}\, A$ for the closure and the closed convex hull of A, and $A \backslash B$ for a set-theoretic difference. The distance of a point x from a set A is $d(x, A)$, while $d_H(A, B)$ indicates the Hausdorff distance between two sets. We write $x + A$ for the set $\{x + y;\ y \in A\}$ and $B(A, \varepsilon)$ $[\bar{B}(A, \varepsilon)]$ for the open [closed] ε-neighborhood around the set A. Given a bounded subset A of a Banach space, its Hausdorff measure of non-compactness [4, p. 41] is

$$\beta(A) = \inf\{r > 0 : A \text{ can be covered with finitely many balls of radius } r\}.$$

The Lebesgue measure of a set $J \subseteq \mathbb{R}$ is meas(J). By a function $f\colon [a, b] \to \mathbb{R}^n$ we shall always mean a Lebesgue measurable function. If f is locally integrable, its Lebesgue set is

$$\mathrm{Leb}(f) = \left\{ t\colon \lim_{\varepsilon \to 0} \varepsilon^{-1} \int_{t-\varepsilon}^{t+\varepsilon} |f(s) - f(t)|\, ds = 0 \right\}.$$

It is well known that the complement of $\mathrm{Leb}(f)$ has measure zero [6, p. 93]. If $f \in \mathscr{L}^2([0, 1]; \mathbb{R}^n)$ satisfies

$$\int_0^1 f(\xi)\, d\xi = \omega, \tag{2.1}$$

then for every $x \in \mathbb{R}^n$ one has the useful identity

$$\int_0^1 |f(\xi) - x|^2\, d\xi = \int_0^1 |f(\xi) - \omega|^2\, d\xi + |x - \omega|^2. \tag{2.2}$$

We denote by $\mathscr{K}_n$ the family of all nonempty compact convex subsets of $\mathbb{R}^n$, endowed with the Hausdorff metric. If $\Omega \in \mathscr{K}_n$, its Chebyshev center $c(\Omega)$ is the unique point $\bar{\omega} \in \Omega$ where the function

$$\varphi_\Omega(x) = \max_{\omega \in \Omega} |\omega - x|$$

attains its global minimum [1, p. 74]. The Chebyshev radius of Ω is then

$$r(\Omega) = \max_{\omega \in \Omega} |\omega - c(\Omega)|.$$

Throughout this paper we shall be concerned with multifunctions $F\colon \mathbb{R}^n \to \mathscr{K}_n$ which are Lipschitz continuous w.r.t. the Hausdorff metric. A (Caratheodory) solution of (1.1) is an absolutely continuous function $t \to x(t)$ which satisfies (1.1) almost everywhere on $[0, T]$. For the basic theory of multifunctions and differential inclusions, our general reference is [1].

3. Likelihood Measures

Let $F\colon \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz continuous multifunction with compact convex values and call $\mathscr{F} \subseteq \mathscr{C}([0, T]; \mathbb{R}^n)$ the family of all Caratheodory solutions of (1.1). To each $u \in \mathscr{F}$ we associate a scalar quantity which roughly measures how many solutions v are located in a small neighborhood of u.

DEFINITION 3.1. The (metric) likelihood of a solution u of (1.1) is the limit

$$L(u) = \lim_{\varepsilon \to 0} \beta\{\dot{v} : v \in \mathscr{F} \cap B(u, \varepsilon)\}, \tag{3.1}$$

where β denotes the Hausdorff measure of non-compactness in the space $\mathscr{L}^2([0, T]; \mathbb{R}^n)$.

Observe that the ball $B(u, \varepsilon)$ refers to the $\mathscr{C}^0$ topology on the space of trajectories, while $L(u)$ measures the non-compactness in $\mathscr{L}^2$ of a set of derivatives. To make good use of the above definition, one now needs a practical formula for evaluating the right-hand side of (3.1). To this purpose, define the function $h: \mathbb{R}^n \times \mathscr{K}_n \to \mathbb{R} \cup \{-\infty\}$ by setting

$$h(\omega, \Omega) = \sup\left\{\left(\int_0^1 |f(\xi) - \omega|^2 \, d\xi\right)^{1/2} : f: [0, 1] \to \Omega, \int_0^1 f(\xi)\, d\xi = \omega\right\}, \tag{3.2}$$

with the understanding that $h(\omega, \Omega) = -\infty$ if $\omega \notin \Omega$. One can interpret $h^2(\omega, \Omega)$ as the maximum variance among all random variables supported inside Ω, whose mean is ω. Clearly, $h(\omega, \Omega) = 0$ iff ω is an extreme point of Ω. Relying on Liapunov's Theorem, when $\omega \in \Omega$ one can prove the existence of a function $f: [0, 1] \to \Omega$ for which the supremum in (3.2) is exactly attained [3, Sect. 16]. Therefore, the function h actually denotes a maximum. Our main result characterizes the likelihood in terms of the function h.

THEOREM 3.2. *The likelihood of a solution u of* (1.1) *is given by*

$$L(u) = \left(\int_0^T h^2(\dot{u}(t), F(u(t)))\, dt\right)^{1/2} = \|h(\dot{u}, F(u))\|_{\mathscr{L}^2}. \tag{3.3}$$

In several cases, the computation of h is rather simple. For each compact convex set Ω, let $c(\Omega)$ and $r(\Omega)$ be its Chebyshev center and its Chebyshev radius, respectively. Define the subset $\Omega^* \subseteq \Omega$ as

$$\Omega^* = \overline{\text{co}}\{\omega \in \Omega : |\omega - c(\Omega)| = r(\Omega)\}. \tag{3.4}$$

EXAMPLE 1. If Ω is the box $[a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$, then $\Omega^* = \Omega$.

EXAMPLE 2. If Ω is a triangle, say $\Omega = \overline{\text{co}}\{\omega_1, \omega_2, \omega_3\}$, then $\Omega^* = \Omega$ iff all of its angles are $\leqslant \pi/2$. If one angle, say at ω_1, is strictly greater than $\pi/2$, then $c(\Omega)=\frac{1}{2}(\omega_2+\omega_3)$, $r(\Omega)=\frac{1}{2}|\omega_3-\omega_2|$, and $\Omega^*=\overline{\mathrm{co}}\{\omega_2, \omega_3\}$.

PROPOSITION 3.3. *Let Ω be a compact convex set in $\mathbb{R}^n$. Then*

$$h(\omega, \Omega) \leqslant \sqrt{r^2(\Omega)-|\omega-c(\Omega)|^2}, \tag{3.5}$$

with equality holding for all $\omega \in \Omega^*$.

Proof. If f: $[0, 1] \to \Omega$ satisfies (2.1), then (2.2) yields

$$\int_0^1 |f(\xi)-\omega|^2 \, d\xi + |\omega - c(\Omega)|^2 = \int_0^1 |f(\xi)-c(\Omega)|^2 \, d\xi \leqslant r^2(\Omega). \tag{3.6}$$

This gives (3.5). If $\omega \in \Omega^*$, there exists a piecewise constant function f: $[0, 1] \to \Omega$ which satisfies (2.1) together with

$$|f(\xi)-c(\Omega)| = r(\Omega) \qquad \forall \xi \in [0, 1].$$

For this f, equality holds in (3.6). Therefore, equality also holds in (3.5).

COROLLARY 3.4. *If all sets $F(x)$ in (1.1) have the property that $F^*(x)=F(x)$, then the likelihood of a trajectory $u \in \mathscr{F}$ is*

$$L(u) = \left(\int_0^T r^2(F(u(t))) \, dt - \int_0^T |\dot{u}(t) - c(F(u(t)))|^2 \, dt \right)^{1/2}. \tag{3.7}$$

4. UPPER SEMICONTINUITY

Before proving Theorem 3.2, we need to establish some basic properties of the integrand function h.

PROPOSITION 4.1. (i) *For every $y \in \mathbb{R}^n$, $h(y+\omega, y+\Omega)=h(\omega, \Omega)$.*

(ii) *If $\Omega_1 \subseteq \Omega_2$, then $h(\omega, \Omega_1) \leqslant h(\omega, \Omega_2)$.*

(iii) *For each fixed $\Omega \in \mathscr{K}_n$, the function $\omega \to h^2(\omega, \Omega)$ is strictly concave down on Ω.*

The first two assertions are obvious. To prove (iii), let $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$, $0<\lambda<1$. Choose f_1, f_2: $[0, 1] \to \Omega$ such that

$$\int_0^1 f_i(\xi) \, d\xi = \omega_i, \qquad \int_0^1 |f_i(\xi)-\omega_i|^2 \, d\xi = h^2(\omega_i, \Omega).$$

If $\omega = \lambda\omega_1 + (1-\lambda)\,\omega_2$ and $f\colon [0, 1] \to \Omega$ is defined by

$$f(\xi) = \begin{cases} f_1\left(\dfrac{\xi}{\lambda}\right) & \text{if} \quad 0 \leqslant \xi \leqslant \lambda, \\ f_2\left(\dfrac{\xi-\lambda}{1-\lambda}\right) & \text{if} \quad \lambda < \xi \leqslant 1, \end{cases}$$

then, using (2.2), one obtains

$$\int_0^1 f(\xi)\,d\xi = \omega,$$

$$\begin{aligned}\int_0^1 |f(\xi)-\omega|^2\,d\xi &= \lambda\int_0^1 |f_1(\xi')-\omega|^2\,d\xi' + (1-\lambda)\int_0^1 |f_2(\xi'')-\omega|^2\,d\xi'' \\ &= \lambda[h^2(\omega_1, \Omega) + |\omega_1 - \omega|^2] \\ &\quad + (1-\lambda)[h^2(\omega_2, \Omega) + |\omega_2 - \omega|^2].\end{aligned}$$

Since $\omega_1, \omega_2 \neq \omega$, this yields

$$h^2(\omega, \Omega) > \lambda h^2(\omega, \Omega) + (1-\lambda)\,h^2(\omega_2, \Omega),$$

proving (iii). Of course, the function $\omega \to h(\omega, \Omega)$ is then concave down as well.

PROPOSITION 4.2. *The function $h\colon \mathbb{R}^n \times \mathscr{K}_n \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous.*

Proof. By (i) and (ii) in Proposition 4.1, it suffices to show that

$$\inf_{\varepsilon > 0} h(0, \bar{B}(\Omega, \varepsilon)) = h(0, \Omega) \qquad \forall \Omega \in \mathscr{K}_n. \tag{4.1}$$

Here 0 denotes the origin in $\mathbb{R}^n$. Let Ω be given. Since the case $0 \notin \Omega$ is trivial, in the following we assume $0 \in \Omega$. Define the vector space

$$V = \{v \in \mathbb{R}^n : \exists \eta > 0 \text{ such that } \eta v \in \Omega,\ -\eta v \in \Omega\} \tag{4.2}$$

and set $\Omega_V = \Omega \cap V$.

Choose radii $R > r > 0$ such that

$$\Omega \subseteq \bar{B}(0, R), \qquad \bar{B}(0, r) \cap V \subseteq \Omega_V \tag*{(4.3)$_{1,2}$}$$

Let any $\varepsilon > 0$ be given. It is not restrictive to assume

$$\varepsilon(R+\varepsilon)/r < 1 - \varepsilon. \tag{4.4}$$

Define $\Omega_\varepsilon = \overline{\text{co}}\{\Omega \backslash B(\Omega_V, \varepsilon)\}$ and observe that $\Omega_V \cap \Omega_\varepsilon = \varnothing$. An auxiliary result is now proved.

LEMMA 4.3. *There exists $\delta \in (0, \varepsilon]$ such that*

$$\text{meas}\{\xi \in [0, 1] : d(f(\xi), \Omega_V) \geqslant \varepsilon\} \leqslant \varepsilon \tag{4.5}$$

for every function f satisfying the conditions

$$f(\xi) \in \bar{B}(\Omega, \delta) \qquad \forall \xi \in [0, 1]; \qquad \int_0^1 f(\xi)\, d\xi = 0. \tag*{$(4.6)_{1,2}$}$$

Proof of the Lemma. Observe first that the function

$$\psi(x) = d\left(\frac{-\varepsilon x}{1-\varepsilon}, \Omega\right) \tag{4.7}$$

is continuous and strictly positive on Ω_ε. By continuity and compactness, there exists $\delta \in (0, \varepsilon]$ such that

$$\min\{\psi(x) : x \in \overline{\text{co}}(\bar{B}(\Omega, \delta) \backslash B(\Omega_V, \varepsilon))\} > \delta. \tag{4.8}$$

Assume now that f satisfies (4.6). Write $[0, 1] = J \cup J^c$, with

$$J = \{\xi : d(f(\xi), \Omega_V) \geqslant \varepsilon\}, \qquad J^c = [0, 1] \backslash J. \tag{4.9}$$

By writing $(4.6)_2$ in the form

$$\text{meas}(J) \cdot \frac{\int_J f(\xi)\, d\xi}{\text{meas}(J)} + (1 - \text{meas}(J)) \cdot \frac{\int_{J^c} f(\xi)\, d\xi}{1 - \text{meas}(J)} = 0$$

we exhibit two vectors

$$\omega_1 \in \overline{\text{co}}(\bar{B}(\Omega, \delta) \backslash B(\Omega_V, \varepsilon)), \qquad \omega_2 \in \bar{B}(\Omega, \delta) \tag{4.10}$$

such that $\text{meas}(J) \cdot \omega_1 + (1 - \text{meas}(J)) \cdot \omega_2 = 0$, i.e.,

$$\omega_2 = \frac{-\text{meas}(J)}{1 - \text{meas}(J)} \cdot \omega_1.$$

If now $\text{meas}(J) > \varepsilon$, we would have

$$\psi(\omega_1) = d\left(\frac{-\varepsilon\omega_1}{1-\varepsilon}, \Omega\right) \leqslant d(\omega_2, \Omega) \leqslant \delta,$$

a contradiction with (4.8). This establishes the lemma.

We now return to the proof of Proposition 4.2. With $\delta > 0$ chosen according to the previous lemma, let f be a function satisfying (4.6) for which

$$\int_0^1 |f(\xi)|^2 \, d\xi = h^2(0, \bar{B}(\Omega, \delta)). \tag{4.11}$$

We will construct a function $\hat{f}$: $[0, 1] \to \Omega_V$ such that

$$\int_0^1 \hat{f}(\xi) \, d\xi = 0 \tag{4.12}$$

and whose $\mathscr{L}^2$ norm is close to the $\mathscr{L}^2$ norm of f. This will provide the needed lower bound on $h(0, \Omega)$. As a preliminary, call π_V the perpendicular projection on the subspace V and define f_V as the composition $\pi_V \circ f$. By (4.5) there exists a set $J \subseteq [0, 1]$ with $\text{meas}(J) \leqslant \varepsilon$ such that

$$d(f(\xi), V) \leqslant d(f(\xi), \Omega_V) \leqslant \varepsilon \qquad \forall \xi \notin J.$$

This implies, in particular,

$$|f_V(\xi)| \geqslant |f(\xi)| - \varepsilon \qquad \forall \xi \notin J. \tag{4.13}$$

Using Liapunov's theorem [7] and recalling (4.4), choose $I \subseteq J^c$ such that

$$\text{meas}(I) = \varepsilon(R + \delta)/r, \tag{4.14}$$

$$\frac{\text{meas}(I)}{\text{meas}(J^c)} \int_{J^c} f_V(\xi) \, d\xi = \int_I f_V(\xi) \, d\xi. \tag{4.15}$$

To construct the function $\hat{f}$, first define

$$\hat{f}(\xi) = \frac{r}{r + \varepsilon} f_V(\xi) \qquad \text{if} \quad \xi \notin I \cup J, \tag{4.16}$$

then set

$$\hat{f}(\xi) = \frac{-1}{\text{meas}(I \cup J)} \int_{(I \cup J)^c} \frac{r}{r + \varepsilon} f_V(\xi) \, d\xi \tag{4.17}$$

for $\xi \in I \cup J$. This constant value of $\hat{f}$ on $I \cup J$ is chosen precisely to satisfy (4.12). Let us check that $\hat{f}(\xi) \in \Omega_V$ for all ξ. If $\xi \notin I \cup J$, one has

$$f_V(\xi) \in V \cap \bar{B}(\Omega_V, \varepsilon),$$

hence $f_V(\xi)=w+w'$ with $w\in\Omega_V$, $w'\in V$, $|w'|\leqslant\varepsilon$. Therefore, recalling $(4.3)_2$,

$$\hat{f}(\xi)=\frac{r}{r+\varepsilon}\,w+\frac{r}{r+\varepsilon}\,w'\in\overline{\mathrm{co}}\left\{w,\frac{r}{\varepsilon}\,w'\right\}\subseteq\Omega_V.$$

On the other hand, if $\xi\in I\cup J$, (4.17) implies $\hat{f}(\xi)\in V$. Moreover, by (4.14), (4.15) and $(4.3)_1$,

$$\begin{aligned}|\hat{f}(\xi)|&\leqslant\frac{1}{\mathrm{meas}(I)}\left|\int_{J^c}f_V(\xi)\,d\xi\right|=\frac{r}{\varepsilon(R+\delta)}\left|\int_J f_V(\xi)\,d\xi\right|\\&\leqslant\frac{r}{\varepsilon(R+\delta)}\cdot\varepsilon(R+\delta)=r.\end{aligned}$$

Because of $(4.3)_2$, we again conclude that $\hat{f}(\xi)\in\Omega_V$.

To estimate the $\mathscr{L}^2$ norm of $\hat{f}$, observe that, by (4.5), (4.14), and $(4.3)_1$,

$$\begin{aligned}&\mathrm{meas}(I\cup J)\leqslant\varepsilon(R+\delta)/r+\varepsilon,\\&(|f(\xi)|-\varepsilon)^2\geqslant|f(\xi)|^2-2\varepsilon(R+\delta).\end{aligned}$$

Using (4.11), (4.13) one thus obtains

$$\begin{aligned}&\int_0^1|\hat{f}(\xi)|^2\,d\xi\\&\quad\geqslant\frac{r}{r+\varepsilon}\int_{(I\cup J)^c}(|f(\xi)|-\varepsilon)^2\,d\xi\\&\quad\geqslant\frac{r}{r+\varepsilon}\int_0^1|f(\xi)|^2\,d\xi-2\varepsilon(R+\delta)-(R+\delta)^2\,\varepsilon(R+r+\delta)/r\\&\quad\geqslant\frac{r}{r+\varepsilon}\,[h^2(0,\bar{B}(\Omega,\delta))]-\varepsilon[(R+\delta)^2\,(R+r+\delta)/r+2(R+\delta)].\end{aligned}\tag{4.18}$$

The above relation shows that, for any $\varepsilon>0$, there exists $\delta\in(0,\varepsilon]$ such that $h^2(0,\Omega)$ is greater or equal than the right-hand side of (4.18). This proves (4.1).

The previous result implies a technical lemma, which we record here for future use.

LEMMA 4.4. *For every $\Omega \in \mathscr{K}_n$, $\omega \in \Omega$, $\varepsilon > 0$, there exists $\eta > 0$ with the following property. If y: $[0, 1] \to \bar{B}(\Omega, \eta)$ satisfies*

$$\left| \int_0^1 y(\xi)\, d\xi - \omega \right| \leqslant \eta, \tag{4.19}$$

then

$$\int_0^1 |y(\xi) - \omega|^2\, d\xi \leqslant h^2(\omega, \Omega) + \varepsilon. \tag{4.20}$$

Indeed,

$$\int_0^1 |h(\xi) - \omega|^2\, d\xi \leqslant \sup\{h^2(\omega', \bar{B}(\Omega, \eta)) : |\omega' - \omega| \leqslant \eta\}.$$

Another useful consequence of Propositions 4.1, 4.2 is:

COROLLARY 4.5. *If u is any solution of (1.1), the function $t \to h(\dot{u}(t), F(u(t)))$ is Lebesgue measurable.*

5. THE UPPER BOUND

Let u be any Caratheodory solution of (1.1). To establish Theorem 3.2, call α the right-hand side of (3.3). By Corollary 4.5, the integral is well defined. In this section we show that α is an upper bound for $L(u)$. In Section 6 we will show that α is also a lower bound, completing the proof.

Let $\varepsilon > 0$ be given and, for simplicity, assume

$$F(x) \subseteq B(0, R) \qquad \forall x \in \mathbb{R}^n \tag{5.1}$$

for some constant R, which is not restrictive.

LEMMA 5.1. *If $t \in [0, T]$ is a Lebesgue point for $\dot{u}$, there exists $\rho_t > 0$ for which the following holds. For each $\rho \in (0, \rho_t]$, there exists $\delta > 0$ such that, if a solution v of (1.1) satisfies*

$$|v(t) - u(t)| < \delta, \qquad |v(t+\rho) - u(t+\rho)| < \delta, \tag{5.2}$$

then

$$\int_t^{t+\rho} |\dot{v}(s) - \dot{u}(t)|^2\, ds \leqslant \rho[h^2(\dot{u}(t), F(u(t))) + \varepsilon]. \tag{5.3}$$

Proof. Setting $\omega = \dot{u}(t)$, $\Omega = F(u(t))$ in Lemma 4.4, choose $\eta \in (0, 1]$ such that (4.19) implies (4.20). Choose $\rho_t > 0$ so small that

$$F(x) \subseteq \bar{B}(F(u(t)), \eta) \qquad \forall x \in \bar{B}(u(t), (R+1)\,\rho_t), \tag{5.4}$$

$$\frac{1}{\rho} \int_t^{t+\rho} |\dot{u}(s) - \dot{u}(t)|\, ds \leqslant \eta/4 \qquad \forall \rho \in (0, \rho_t]. \tag{5.5}$$

We claim that with the above choices the conclusion of the lemma holds. Indeed, if $0 < \rho < \rho_t$, set $\delta = \rho\eta/4$ and let v be any solution of (1.1) satisfying (5.2). By (5.4) we have

$$\dot{v}(s) \in \bar{B}(F(u(t)), \eta) \qquad \forall s \in [t, t+\rho]. \tag{5.6}$$

After the change of variable $\xi = (s-t)/\rho$, (5.2) and (5.5) together yield

$$\left| \int_0^1 \dot{v}(\xi)\, d\xi - \dot{u}(t) \right| \leqslant 3\eta/4. \tag{5.7}$$

Therefore, by (4.20),

$$\int_0^1 |\dot{v}(\xi) - \dot{u}(t)|^2\, d\xi \leqslant h^2(\dot{u}(t), F(u(t))) + \varepsilon. \tag{5.8}$$

Written in terms of the original variable s, this is precisely (5.3).

Returning to the main proof, for each $t \in \text{Leb}(\dot{u})$ choose $\rho_t > 0$ according to Lemma 5.1. If t also lies in the Lebesgue set of the measurable map $s \to h(\dot{u}(s), F(u(s)))$, by possibly shrinking the value of ρ_t we can assume that

$$\int_t^{t+\rho} |h^2(\dot{u}(s), F(u(s))) - h^2(\dot{u}(t), F(u(t)))|\, ds \leqslant \rho\varepsilon \tag{5.9}$$

for each $\rho \in (0, \rho_t]$. The family of all closed intervals $[t, t+\rho]$ with $0 < \rho \leqslant \rho_t$ is a Vitali covering of $\text{Leb}(\dot{u}) \cap \text{Leb}(h(\dot{u}, F(u)))$. Since these Lebesgue sets have full measure in $[0, T]$, using Vitali's theorem [11, p. 109] one can select finitely many disjoint intervals $J_i = [t_i, t_i + \rho_i]$ and corresponding $\delta_i > 0$, $i = 1, ..., N$, for which (5.3) holds, together with

$$\text{meas}\left([0, T] \Big\backslash \bigcup_{i=1}^{N} J_i \right) < \varepsilon. \tag{5.10}$$

Set $\delta = \min\{\delta_i;\, i = 1, ..., N\}$ and define the function $y \in \mathscr{L}^2$:

$$y(t) = \dot{u}(t_i) \qquad \text{if} \quad t \in J_i,$$

$$y(t) = 0 \qquad \text{if} \quad t \notin \bigcup J_i.$$

We now show that the set $\{\dot{v}; v \in \mathscr{F} \cap B(u, \delta)\}$ can be covered with just one ball centered at y, whose radius is very close to α. Indeed, let v be any solution of (1.1) contained in $B(u, \delta)$. Calling $J = \bigcup J_i$, $J^c = [0, T] \backslash J$ and using (5.3), (5.9), (5.10) one obtains

$$\begin{aligned}\int_0^T |\dot{v}(t) - y(t)|^2 \, dt &\leqslant \sum_{i=1}^N \int_{J_i} |\dot{v}(t) - \dot{u}(t_i)|^2 \, dt + 4R^2 \operatorname{meas}(J^c) \\ &\leqslant \sum_{i=1}^N \rho_i [h^2(\dot{u}(t_i), F(u(t_i))) + \varepsilon] + 4R^2 \varepsilon \\ &\leqslant \int_J h^2(u(s), F(u(s))) \, ds + 2\varepsilon \operatorname{meas}(J) + 4R^2 \varepsilon \\ &\leqslant \int_0^T h^2(\dot{u}(s), F(u(s))) \, ds + \varepsilon(2T + 4R^2). \qquad (5.11)\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the inequality $L(u) \leqslant \alpha$ is proved.

6. The Lower Bound

The proof that the right-hand side of (3.3) is a lower bound for $L(u)$ is achieved after a series of lemmas.

Lemma 6.1. *Let $\Omega \in \mathscr{K}_n$, $\omega \in \Omega$, $\tau > 0$, and define the set of derivatives*

$$E = \{\dot{v} : v \text{ is absolutely continuous, } v(0) = 0, v(\tau) = \tau\omega, \dot{v} \in \Omega \text{ a.e. in } [0, \tau]\} \subseteq \mathscr{L}^2([0, \tau]; \mathbb{R}^n).$$

Then

$$\beta_{[0,\tau]}(E) = \left(\int_0^\tau h^2(\omega, \Omega) \, dt \right)^{1/2} = \sqrt{\tau}\, h(\omega, \Omega), \qquad (6.1)$$

where $\beta_{[0,\tau]}$ denotes the Hausdorff measure of non-compactness on $\mathscr{L}^2([0, \tau]; \mathbb{R}^n)$.

Proof. The definition of h implies that E is contained in the single closed ball $B \subseteq \mathscr{L}^2$ having the constant function $y(t) = \omega$ as center, with radius $\sqrt{\tau}\, h(\omega, \Omega)$. Therefore, $\beta_{[0,\tau]}(E)$ is not greater than the right-hand side of (6.1). To show the converse inequality, choose a function $\varphi: [0, 1] \to \Omega$ such that

$$\int_0^1 \varphi(\xi) \, d\xi = \omega, \qquad \int_0^1 |\varphi(\xi) - \omega|^2 \, d\xi = h^2(\omega, \Omega). \qquad (6.2)$$

Extend φ to the whole real line by periodicity (with period 1), and define the sequence of functions φ_ν: $[0, 1] \to \Omega$ by setting

$$\varphi_\nu(\xi) = \varphi(\nu\xi). \tag{6.3}$$

If now ψ: $[0, \tau] \to \mathbb{R}^n$ is a constant function, one trivially has

$$\begin{aligned}\lim_{\nu \to \infty} \int_0^\tau \left|\varphi_\nu\left(\frac{s}{\tau}\right) - \psi(s)\right|^2 ds &= \int_0^\tau \left|\varphi_\nu\left(\frac{s}{\tau}\right) - \omega\right|^2 ds + \int_0^\tau |\omega - \psi(s)|^2\, ds \\ &\geqslant \int_0^\tau \left|\varphi\left(\frac{s}{\tau}\right) - \omega\right|^2 ds = \tau[h^2(\omega, \Omega)]. \end{aligned} \tag{6.4}$$

Observe that (6.4) is still valid if ψ is any piecewise constant function. A standard approximation technique then shows that (6.4) continues to hold for every $\psi \in \mathscr{L}^2$. Given any $\varepsilon > 0$, if $\{y_1, ..., y_N\}$ is any finite subset of $\mathscr{L}^2([0, \tau]; \mathbb{R}^n)$, for ν large enough we have

$$\int_0^\tau \left|\varphi_\nu\left(\frac{s}{\tau}\right) - y_i(s)\right|^2 ds > \tau[h^2(\omega, \Omega) - \varepsilon] \tag{6.5}$$

for all i. Therefore, the N balls $B(y_i, \sqrt{\tau}\,[h^2(\omega, \Omega) - \varepsilon]^{1/2})$ cannot cover E. Since $\varepsilon > 0$ was arbitrary, we conclude that $\beta_{[0,\tau]}(E) \geqslant \sqrt{\tau}\, h(\omega, \Omega)$.

In the following we still assume (5.1) and call λ a Lipschitz constant for F. Fix any solution u of (1.1) and any integer $m \geqslant 1$. Divide the interval $[0, T]$ into m equal subintervals, inserting the points $t_k = kT/m$. Since R is a Lipschitz constant for u, one has

$$F(u(t)) \subseteq \bar{B}(F(u(t_{k-1})), \lambda RT/m) \qquad \forall t \in [t_{k-1}, t_k]. \tag{6.6}$$

Define the set

$$\begin{aligned} E_m = \{\dot{v}\colon v \text{ is absolutely continuous and, for all } k,\ v(t_k) = u(t_k), \\ \dot{v}(t) \in \bar{B}(F(u(t_{k-1})), \lambda RT/m) \text{ a.e. in } [t_{k-1}, t_k]\}. \end{aligned}$$

LEMMA 6.2. *The Hausdorff measure of non-compactness of the set E_m satisfies*

$$\beta(E_m) \geqslant \left(\int_0^T h^2(\dot{u}(t), F(u(t)))\, dt\right)^{1/2}. \tag{6.7}$$

Proof. For each $k=1, \dots, m$, (6.6) and the concavity of h^2 imply

$$\int_0^T h^2(\dot{u}(t), F(u(t)))\,dt \leqslant \sum_{k=1}^{m} \int_{t_{k-1}}^{t_k} h^2(\dot{u}(t), \bar{B}(F(u(t_{k-1})), \lambda RT/m))\,dt$$

$$\leqslant \frac{T}{m} \sum_{k=1}^{m} h^2(\omega_k, \Omega_k), \tag{6.8}$$

where

$$\omega_k = \frac{m}{T}(u(t_k) - u(t_{j-1})), \qquad \Omega_k = \bar{B}(F(u(t_{k-1})), \lambda RT/m).$$

It therefore suffices to show that

$$\beta^2(E_m) = \frac{T}{m} \sum_{k=1}^{m} h^2(\omega_k, \Omega_k). \tag{6.9}$$

The proof of (6.9) is achieved by applying the same technique of Lemma 6.1 to each subinterval $[t_{k-1}, t_k]$. Fix $\varepsilon > 0$. For each $k=1, \dots, m$, choose a function $\varphi_k: [0, 1] \to \Omega_k$ such that

$$\int_0^1 \varphi_k(\xi)\,d\xi = \omega_k, \qquad \int_0^1 |\varphi_k(\xi) - \omega_k|^2\,d\xi = h^2(\omega_k, \Omega_k), \tag{6.10}$$

and extend φ_k to the whole real line, by periodicity. Define a sequence of functions $g_\nu: [0, T] \to \bigcup \Omega_k$ by setting

$$g_\nu(s) = \varphi_k\left(\frac{\nu m}{T}(s - t_{k-1})\right) \qquad s \in (t_{k-1}, t_k]. \tag{6.11}$$

As in the previous lemma, for any function $\psi \in \mathscr{L}^2$ and any $\varepsilon > 0$, (6.10) implies

$$\lim_{\nu \to \infty} \int_0^T |g_\nu(s) - \psi(s)|^2\,ds = \sum_{k=1}^{m} \int_{t_{k-1}}^{t_k} (|g_\nu(s) - \omega_k|^2 + |\psi(s) - \omega_k|^2)\,ds$$

$$> \frac{T}{m} \sum_{k=1}^{n} [h^2(\omega_k, \Omega_k) - \varepsilon]. \tag{6.12}$$

Since $\varepsilon > 0$ and ψ can be chosen arbitrarily, (6.12) shows that E_m cannot be covered in $\mathscr{L}^2$ with finitely many balls whose radius is smaller than

$$\bar{r} \doteq \left[(T/m) \sum_{k=1}^{m} h^2(\omega_k, \Omega_k) \right]^{1/2}.$$

Therefore, $\beta^2(E_m)$ is greater or equal than the right-hand side of (6.9). The opposite inequality is trivial, since E_m is contained in the single closed ball with radius $\bar{r}$, having as center the piecewise constant function

$$y(s) = \omega_k \qquad s \in (t_{k-1}, t_k].$$

The previous lemma provides a lower bound on the measure of non-compactness for a set E_m of derivatives of approximate solutions of (1.1). By a classical theorem of Filippov [5], for each approximate solution v there exists a nearby genuine solution, say w. A careful estimate of the $\mathscr{L}^2$ distance $\|\dot{v} - \dot{w}\|$, combined with (6.7), will yield the lower bound on $L(u)$.

LEMMA 6.3. *Let Φ be an arbitrary map from E_m into $\mathscr{L}^2$. If*

$$\|\Phi(y) - y\| \leqslant \sigma \qquad \forall y \in E_m \tag{6.13}$$

then the measure of non-compactness of the image $\Phi(E_m)$ satisfies

$$\beta(\Phi(E_m)) \geqslant \beta(E_m) - \sigma. \tag{6.14}$$

Indeed, if $\Phi(E_m)$ is covered with finitely many balls $B(y_i, \rho)$ for some radius ρ, then the balls $B(y_i, \rho + \sigma)$ cover E_m.

LEMMA 6.4. *For each $m \geqslant 1$, setting*

$$\eta_m = m^{-1}[2\lambda RT(e^{\lambda t} - 1) + RT], \tag{6.15}$$

with the same notation as (3.1) *one has*

$$\beta\{\dot{v} : v \in \mathscr{F} \cap B(u, \eta_m)\} \geqslant \beta(E_m) - \frac{2\lambda RT^{3/2}e^{\lambda T}}{m}. \tag{6.16}$$

Proof. Let v be an absolutely continuous function such that, for all k,

$$v(t_k) = u(t_k), \qquad \dot{v}(t) \in \bar{B}(F(u(t_{k-1})), \lambda RT/m)) \qquad \text{a.e. in } [t_{k-1}, t_k].$$

Observe that

$$d(\dot{v}(t), F(v(t))) \leqslant d_H(F(v(t)), \bar{B}(F(u(t_{k-1})), \lambda RT/m)) \leqslant 2\lambda RT/m$$

for almost every $t \in (t_{k-1}, t_k]$. By Theorem 1 in [5], close to the approximate solution v there exists a genuine solution w of (1.1), satisfying

$$|w(t) - v(t)| \leqslant \frac{2RT}{m}(e^{\lambda t} - 1), \tag{6.17}$$

$$|\dot{w}(t) - \dot{v}(t)| \leqslant \frac{2\lambda RT}{m} e^{\lambda t} \qquad \text{a.e. in } [0, T]. \tag{6.18}$$

By (6.2), $|u(t)-v(t)| \leqslant RT/m$ for all t. Therefore, $w \in \mathscr{F} \cap \bar{B}(u, \eta_m)$. Using the axiom of choice, one can construct a function $\Phi: \dot{v} \to \dot{w}$ defined on E_m, such that, by (6.18),

$$\|\dot{v} - \Phi(\dot{v})\|_{\mathscr{L}^\infty} \leqslant \frac{2\lambda RTe^{\lambda T}}{m}. \tag{6.19}$$

In the $\mathscr{L}^2$ norm, (6.19) becomes

$$\|\dot{v} - \Phi(\dot{v})\|_{\mathscr{L}^2} \leqslant \frac{2\lambda RTe^{\lambda T}\sqrt{T}}{m}.$$

An application of Lemma 6.3 now yields (6.16).

Finally, putting together (6.7) with (6.16) we obtain

$$\begin{aligned}\lim_{\varepsilon \to 0} \beta\{\dot{v}: v \in \mathscr{F} \cap B(u, \varepsilon)\} &= \lim_{m \to \infty} \beta\{\dot{v}: v \in \mathscr{F} \cap \bar{B}(u, \eta_m)\} \\ &\geqslant \liminf_{m \to 0} \left[\beta(E_m) - \frac{2\lambda RT^{3/2}e^{\lambda T}}{m}\right] \\ &\geqslant \left(\int_0^T h^2(\dot{u}(t), F(u(t)))\, dt\right)^{1/2},\end{aligned}$$

completing the proof of Theorem 3.2.

By a theorem of Olech [9], the upper semicontinuity of the integrand function h^2 and its concavity w.r.t. ω imply:

COROLLARY 6.5. *The functional $u \to L(u)$ is upper semicontinuous.*

7. The Cauchy Problem

Fix $x_0 \in \mathbb{R}^n$ and consider the family $\mathscr{F}_0$ of all solutions of the Cauchy problem

$$\dot{x}(t) \in F(x(t)), \qquad x(0) = x_0, \qquad t \in [0, T]. \tag{7.1}$$

If $u \in \mathscr{F}_0$, one can again define the likelihood of u as

$$L(u) = \lim_{\varepsilon \to 0} \beta\{\dot{v}: v \in \mathscr{F}_0 \cap B(u, \varepsilon)\}. \tag{7.2}$$

The proof of Theorem 3.2 shows that even in this case our integral formula continues to hold. We now introduce a scalar quantity which roughly measures how many trajectories of (7.1) approach a given point $\bar{x}$ at time T.

DEFINITION 7.1. *The likelihood for a point $\bar{x}$ to be approached by solutions of* (7.1) *at time T is*

$$L(\bar{x}) = \lim_{\varepsilon \to 0} \beta\{\dot{u} : u \in \mathscr{F}_0, |u(T) - \bar{x}| < \varepsilon\}. \tag{7.3}$$

Here, as usual, β denotes the Hausdorff measure of non-compactness in $\mathscr{L}^2([0, T]; \mathbb{R}^n)$. The next result, combined with Theorem 3.2, shows that $L(\bar{x})$ is the value function of a corresponding variational problem.

THEOREM 7.2.

$$L(\bar{x}) = \max\{L(v) : v \in \mathscr{F}_0, v(T) = \bar{x}\}. \tag{7.4}$$

Proof. Call $\mathscr{F}_0^{\bar{x}}$ the set of solutions v of (7.1) which exactly reach $\bar{x}$ at time T, and call α the right-hand side of (7.4). Observe that α is well defined, being the maximum of an upper semicontinuous functional on a compact set. For every $v \in \mathscr{F}_0^{\bar{x}}$ and every $\varepsilon > 0$ one has

$$\beta\{\dot{u} : u \in \mathscr{F}_0, |u(T) - \bar{x}| < \varepsilon\} \geqslant \beta\{\dot{u} : u \in \mathscr{F}_0 \cap B(v, \varepsilon)\}.$$

Letting $\varepsilon \to 0$ we deduce $L(\bar{x}) \geqslant \alpha$. To prove the converse inequality, fix any constant $\alpha' > \alpha$. For each $v \in \mathscr{F}_0^{\bar{x}}$ there exists $\delta > 0$ such that

$$\beta\{\dot{w} : w \in \mathscr{F}_0 \cap B(v, \delta)\} < \alpha'. \tag{7.5}$$

Cover the compact set $\mathscr{F}_0^{\bar{x}}$ with finitely many open balls $B_i = B(v_i, \delta_i)$ for which (7.5) holds. We claim that there exists $\varepsilon > 0$ such that the set

$$A_\varepsilon = \{u \in \mathscr{F}_0 : |u(T) - \bar{x}| < \varepsilon\}$$

is contained in the union of the balls B_i. If not, there would exist a sequence $u_n \in \mathscr{F}_0$ such that $u_n(T) \to \bar{x}$ and

$$u_n \notin \bigcup_i B_i \qquad \forall n \geqslant 1. \tag{7.6}$$

By compactness, a subsequence $u_{n'}$ would converge to some trajectory $\hat{u} \in \mathscr{F}_0^{\bar{x}}$. Since $\hat{u} \in B_j$ for some j, (7.6) yields a contradiction. By our claim, we can now write

$$L(\bar{x}) \leqslant \beta(A_\varepsilon) \leqslant \max_i \beta(\mathscr{F}_0 \cap B(v_i, \delta_i)) < \alpha',$$

proving the theorem.

EXAMPLE 3. In general,

$$L(\bar{x}) \neq \beta\{\dot{u}: u \in \mathscr{F}_0,\ u(T) = \bar{x}\}.$$

Indeed, consider the Cauchy problem in $\mathbb{R}^n$:

$$(\dot{x}_1, \dot{x}_2) \in [-1, 1] \times \{x_1^2\}, \qquad (x_1, x_2)(0) = (0, 0).$$

Set $T = 1$, $\bar{x} = (0, 0)$ and let u_0 be the constant trajectory: $u_0(t) \equiv (0, 0)$ for all t. Then

$$L(\bar{x}) = L(u_0) = 1 \neq \beta\{\dot{u}: u \in \mathscr{F}_0,\ u(1) = \bar{x}\} = \beta\{\dot{u}_0\} = 0.$$

8. PREDICTION AND FILTERING

With the theory developed thus far, it is possible to give a precise mathematical meaning to a class of estimation, prediction, and filtering problems, in a context which is entirely independent of probability theory.

1. PREDICTION PROBLEM. *Given the Cauchy problem* (7.1), *predict the most likely path.*

With the notation of the previous sections, this amounts to finding a trajectory $u \in \mathscr{F}_0$ which maximizes the functional $L(u)$.

2. FILTERING PROBLEM. *Assume that, for the Cauchy problem* (7.1), *a function $y(t) = g(x(t))$ is observed. For each $t \in [0, T]$, give an estimate for $x(t)$, knowing the values $y(s)$ for $0 \leqslant s \leqslant t$.*

For every t, the above problem can be formulated as

$$\max \int_0^t h^2(\dot{u}(s), F(u(s)))\, ds$$

subject to $u \in \mathscr{F}_0$, $g(u(s)) = y(s)$, $0 \leqslant s \leqslant t$.

3. ESTIMATION PROBLEM. *For each θ in a compact set $\boldsymbol{\Theta}$ of parameters, let a Lipschitz continuous multifunction F^θ be given, depending continuously on θ. If some trajectory $u(\cdot)$ of a differential inclusion*

$$\dot{x}(t) \in F^\theta(x(t)), \qquad x(0) = x_0$$

is observed, give an estimate on θ.

This leads to the problem of maximizing the likelihood

$$L^{\theta}(u)=\left(\int_0^T h^2(\dot{u}, F^{\theta}(u(t)))\,dt\right)^{1/2}$$

for u fixed, over all $\theta\in\Theta$.

We remark that each of the above problems involves the computation of the maximum of an upper semicontinuous functional over a compact set. The existence of at least one solution therefore follows from classical theorems [3].

EXAMPLE 4. Consider a system consisting of a signal z and an observation y, satisfying the differential inclusion

$$\dot{z}(t)\in\bar{B}(f(z(t)), r_1) \qquad z(0)=\bar{z}\in\mathbb{R}^m, \tag{8.1}$$

$$\dot{y}(t)\in\bar{B}(g(z(t)), r_2) \qquad z(0)=\bar{z}\in\mathbb{R}^n. \tag{8.2}$$

We regard (8.1), (8.2) as a unique differential inclusion for $x=(z, y)$ in the product space $\mathbb{R}^m\times\mathbb{R}^n$. By Corollary 3.4, the likelihood of a solution $x(\cdot)=(z(\cdot), y(\cdot))$ up to time t is

$$L_t(x)=\left(\int_0^t [(r_1^2+r_2^2)-|\dot{z}(s)-f(z(s))|^2-|\dot{y}(s)-g(z(s))|^2]\,ds\right)^{1/2}.$$

If an absolutely continuous observation $y_0(s)$ is given for $s\in[0, t]$, an optimal estimate for the signal z is thus a function $\hat{z}$ which minimizes the functional

$$J_t(z)=\int_0^t |\dot{z}(s)-f(z(s))|^2+|\dot{y}_0(s)-g(z(s))|^2\,ds, \tag{8.3}$$

subject to the constraints

$$|\dot{z}(s)-f(z(s))|\leqslant r_1, \qquad |\dot{y}_0(s)-g(z(s))|\leqslant r_2.$$

It is interesting to observe the close analogy between the above and the stochastic filtering problem

$$dz(t)=f(z(t))\,dt+dw_n(t) \qquad z(0)=z_0,$$

$$dy(t)=g(z(t))\,dt+dw_m(t) \qquad y(0)=y_0,$$

where w_n, w_m are n- and m-dimensional, independent Brownian motions. Indeed, the (stochastic) maximum likelihood estimator for z, given an absolutely continuous observation $y_0(s)$, $0\leqslant s\leqslant t$, is precisely the one which minimizes the integral (8.3), without any constraints [8]. When f and g are linear, the optimal estimate is given by Kalman's filter.

REFERENCES

1. J. P. AUBIN AND A. CELLINA, "Differential Inclusions," Springer-Verlag, Berlin/New York, 1984.
2. A. CELLINA AND C. MARICONDA, Kuratowski's index of a decomposable set, *Bull. Polish Acad. Sci.*, to appear.
3. L. CESARI, "Optimization—Theory and Applications," Springer-Verlag, Berlin/New York, 1983.
4. K. DEIMLING, "Nonlinear Functional Analysis," Springer-Verlag, Berlin/New York, 1985.
5. A. F. FILIPPOV, Classical solutions of differential equations with multivalued right-hand side, *SIAM J. Control* **5** (1967), 609–621.
6. G. B. FOLLAND, "Real Analysis," Wiley, New York, 1984.
7. P. HALMOS, The range of a vector measure, *Bull. Amer. Math. Soc.* **54** (1948), 416–421.
8. R. E. MORTENSEN, Maximum-likelihood recursive nonlinear filtering, *J. Optim. Theory Appl.* **2** (1968), 386–394.
9. C. OLECH, Weak lower semicontinuity of integral functionals, *J. Optim. Theory Appl.* **19** (1976), 3–16.
10. J. OXTOBY, "Measure and Category," Springer-Verlag, Berlin/New York, 1971.
11. S. SAKS, "Theory of the Integral," Dover, New York, 1964.